

日 本 国 特 許 庁
JAPAN PATENT OFFICE

23.06.03

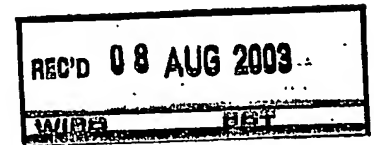
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 2 年 1 1 月 1 日
Date of Application:

出 願 番 号 特 願 2 0 0 2 - 3 1 9 7 5 6
Application Number:
[ST. 10/C] : [J P 2 0 0 2 - 3 1 9 7 5 6]

出 願 人 松下電器産業株式会社
Applicant(s):

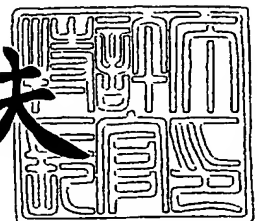


**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

2 0 0 3 年 7 月 2 5 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



BEST AVAILABLE COPY

【書類名】 特許願

【整理番号】 2054041310

【提出日】 平成14年11月 1日

【あて先】 特許庁長官殿

【国際特許分類】 H04N 5/91
G11B 20/12

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 下田代 雅文

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 桑野 裕康

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 酒井 啓行

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 小林 正明

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 松井 謙二

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 メタデータ制作装置及び検索装置

【特許請求の範囲】

【請求項1】 コンテンツに関連するメタデータの制作装置であって、

制作されたコンテンツに合わせて制作されたシナリオ、或いは前記コンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と、

前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、

前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段とを具備することを特徴とするメタデータ制作装置。

【請求項2】 前記音声認識手段は、ジャンル別辞書を複数用意し、コンテンツに適合したジャンルの辞書を選択することを特徴とする請求項1記載のメタデータ制作装置。

【請求項3】 前記音声認識手段は、ジャンル別辞書を複数用意し、コンテンツに適合したジャンルの辞書を選択し、さらに、前記シナリオ、或いはコンテンツの内容から抽出したキーワードを優先認識するようにしたことを特徴とする請求項1または2に記載のメタデータ制作装置。

【請求項4】 コンテンツに関連するメタデータの制作装置であって、

制作されたコンテンツにナレーションの音声信号を入力し、前記音声信号をデータ化する入力手段と、

ジャンル別辞書を複数用意し、コンテンツに適合したジャンルの辞書を選択するとともに、前記コンテンツのシナリオ、或いは前記コンテンツの内容から抽出したキーワードを優先認識するように設定し、前記入力装置でデータ化された音声信号データから、キーワードを認識する音声認識手段と、

前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共に、メタデータファイルに記憶するファイル処理手段とを具備することを特徴とするメタデータ制作装置。

【請求項5】コンテンツに関連するメタデータの検索装置であって、

制作されたコンテンツに合わせて制作されたシナリオ、或いは前記コンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と、

前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、

前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段と、

前記コンテンツファイルの記録位置と前記メタデータファイルの関係を管理する制御ファイルを発生させるコンテンツ情報ファイル処理手段と、

前記コンテンツファイルと、前記メタデータファイルと、前記制御ファイルとを記録する記録手段とを具備したことを特徴とするメタデータ検索装置。

【請求項6】前記コンテンツ情報ファイル処理手段から出力される制御ファイルは、コンテンツの記録時間に合わせた記録手段の前記コンテンツ記録位置を明記するテーブルとし、前記タイムコードから前記コンテンツの記録位置を検索できるようにしたことを特徴とする請求項5記載のメタデータ検索装置。

【請求項7】前記音声認識手段は、ジャンル別辞書を複数用意し、コンテンツに適合したジャンルの辞書を選択し、さらに、前記シナリオ、或いは前記コンテンツの内容から抽出したキーワードを優先認識するようにしたことを特徴とする請求項5または6に記載のメタデータ検索装置。

【請求項8】コンテンツに関連するメタデータの検索装置であって、

制作されたコンテンツに合わせて制作されたシナリオ、或いは前記コンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と、

ジャンル別辞書を複数用意し、コンテンツに適合したジャンルの辞書を選択し、前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、

前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信

号との時間位置を示すタイムコード共にメタデータファイルに記憶するファイル処理手段と、

前記コンテンツファイルの記録位置と前記メタデータファイルの関係を管理する制御ファイルを発生させるコンテンツ情報ファイル処理手段と、

前記コンテンツファイルと、前記メタデータファイルと、前記制御ファイルとを一緒に記録する記録手段と、

前記音声認識手段で用いた共通辞書からキーワードを選定し、前記選定したキーワードが記録されている前記メタデータファイルを特定し、前記記録手段に記録されているコンテンツの中から、検索したいコンテンツを検索し、前記制御ファイルから検索したいシーンの記録位置を検索する検索手段とを具備し、

コンテンツ管理用メタデータを自動発生し、前記メタデータから前記記録手段に記録されたコンテンツファイルの記録位置を特定することを特徴とするメタデータ制作装置及び検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、コンテンツ制作におけるメタデータ制作装置及び検索装置に関するものである。

【0002】

【従来の技術】

近年、映像・音声コンテンツの制作において、これらコンテンツに関連したメタデータの付与することがおこなわれている。

【0003】

しかしながら、上記メタデータの付与は、制作された映像・音声コンテンツのシナリオあるいはナレーション原稿をもとに、制作された映像・音声コンテンツを再生しながらメタデータとすべき情報を確認し、手作業でコンピュータ入力することにより制作する方法が一般的であり、相当な労力の必要な方法であった。

【0004】

また、カメラ撮影時に音声認識を用いタグ付けをするシステムは存在するが、

撮影と同時に使用されるものに過ぎなかった。(特許文献1参照)

【0005】

【特許文献1】

特開平09-130736号公報

【0006】

【発明が解決しようとする課題】

本願発明は、上記従来の問題点に係る課題を解決することを目的とするものであって、制作された映像・音声コンテンツを再生することによりメタデータとすべき情報を確認し、音声入力でコンピュータ等に入力することにより前記メタデータを制作し、検索するシステムを提供することを目的とする。

【0007】

【課題を解決するための手段】

上記課題を解決するために本願発明は、制作されたコンテンツに合わせて制作されたシナリオ、或いは前記コンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段とを備えたものである。

これにより、従来キーボードで入力し、制作していたメタデータを、音声認識を用いて音声入力し、自動的にタイムコード付きのメタデータを制作することが可能となる。

特に、数秒単位の間隔でメタデータを付与する場合は、キー入力では困難であるが、本構成によれば、数秒単位間隔であっても効率よく、メタデータを付与できる。

また、音声入力する際に、事前シナリオ、或いはコンテンツの内容からキーワードを限定して入力する構成となっているので、音声認識の認識率を改善できる。

【0008】

【発明の実施の形態】

本発明の請求項 1 から 4 に係る発明は、制作されたコンテンツに合わせて制作されたシナリオから抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、前記音声認識手段から出力されたキーワードをコンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段とを具備したことを特徴とするメタデータ制作装置である。

【0009】

本発明の請求項 5 から 7 に係る発明は制作されたコンテンツに合わせて制作されたシナリオ、或いはコンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段と、前記コンテンツファイルと前記メタデータファイルの関係を管理する制御ファイルを発生させるコンテンツ情報ファイル処理手段と、前記コンテンツファイルと、前記メタデータファイルと、前記制御ファイルとを記録する記録手段とを具備したことを特徴とするメタデータ検索装置である。

【0010】

本発明の請求項 8 に係る発明は制作されたコンテンツに合わせて制作されたシナリオ、或いはコンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と、ジャンル別辞書を複数用意し、コンテンツに適合したジャンルの辞書を選択し、前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段と、前記コンテンツファイルの記録位置と前記メタデータファイルの関係を管理する制御ファイルを発生させるコンテンツ情報ファイル処理手段と、前記コンテンツファイルと、前記メタデータファイルと、前記制御ファイルとを一緒に

記録する記録手段と、検索したいコンテンツの分野に適合した、前記音声認識手段で用いた共通辞書からキーワードを選定し、前記選定したキーワードが記録されている前記メタデータファイルを特定し、前記記録手段に記録されているコンテンツの中から、検索したいコンテンツを検索し、前記制御ファイルから検索したいシーンの記録位置を検索する検索手段とを具備し、前記メタデータから前記記録手段に記録されたコンテンツファイルの記録位置とを特定することを特徴とするメタデータ検索装置である。

【0011】

以下、本発明の実施の形態について図面を用いて説明する。

(実施の形態1)

図1は、本発明の実施の形態1によるメタデータ制作装置の構成を示すブロック図である。

【0012】

図1において、1はコンテンツデータベース(DB)、2は入力手段、3は音声認識手段、4は辞書データベース(DB)、5はファイル処理手段、11は映像モニタである。

コンテンツDB1は、例えばVTR（あるいはハードディスクで構成された映像・音声信号再生手段、あるいは半導体メモリなどのメモリ手段を記録媒体とする映像・音声信号再生手段、あるいは光学記録式または磁気記録式などの回転型ディスクで構成された映像・音声信号再生手段、更には伝送されてきたあるいは放送されてきた映像・音声信号を1次記録し、再生する映像・音声再生手段などの、コンテンツに合わせたタイムコードを発生しながら再生する手段を備えたコンテンツ記録手段）である。

コンテンツDB1から再生されたタイムコード付き、映像信号は映像モニタ11に出力され、前記映像モニタ11で映出される。

次に、前記映像モニタ11に映出されたコンテンツに合わせて、ナレータ12がマイクロホーンを用いてナレーションの音声信号を入力する。この際、ナレータは映像モニタ11に映しだされたコンテンツ、あるいは、タイムコードを確認し、シナリオ、或いはナレーション原稿、或いはコンテンツの内容などを基に抽出

されたコンテンツ管理用キーワードを発声し、マイクロホーンを用いてナレーションとして音声信号を入力する。

従って、前記したように入力される音声信号を前もってシナリオ等から限定されたキーワードを使用することによって、後段の音声認識手段3での認識率を改善させることができる。

次に、入力手段2ではマイクロホーンから入力された音声信号を、コンテンツDB1から出力されている垂直同期信号に同期したクロックで、前記音声信号をデータ化する。

次に、入力手段2でデータ化された音声信号データは、音声認識手段3に入力される。また、同時に、音声認識に必要な辞書が辞書DB4から供給される。

ここで、使用する音声認識用辞書を端子102から辞書DB4に設定する。

例えば、図2に示すように各分野別に辞書DB4が構成されていたとすると、使用する分野を端子102（例えば、キー入力できるキーボード端子）から設定する。

料理番組の場合は、料理—日本料理—料理法—野菜炒め等を端子102から辞書DB4を設定する。

前記のように辞書DB4を設定することで使用する単語、および、音声認識すべき単語を制限し、音声認識手段3の認識率を改善する。

また、更に、図1にもどり、端子102からシナリオ、あるいは、シナリオ原稿、あるいはコンテンツの内容から抽出されたキーワードを入力する。

例えば、料理番組の場合は、図3に示すレシピを端子102から入力する。

従って、レシピに記入されている単語が音声信号として入力されてくる可能性が高いので、辞書DB4では端子102から入力されたレシピ単語の認識優先度を明示し、優先して音声認識を行うようにする。

例えば、「柿」と「貝のカキ」が辞書中にあった場合、端子102から入力されたレシピ単語が「貝のカキ」のみの場合は、「貝のカキ」に優先順位1がつけられる。

音声認識手段3では、「かき」という音声を認識した場合、辞書DB4に設定された単語の優先順位1が明記されている「貝のカキ」と認識する。

従って、辞書DB4では、端子102から入力される分野で単語を限定し、更に、シナリオを端子102から入力して単語の優先度を明示することで、音声認識手段3での認識率を改善させることができる。

図1にもどり、音声認識手段3では、辞書DB4から供給された辞書に従って、入力手段2から入力された音声信号データを認識し、メタデータを生成する。次に、音声認識手段3から出力されたメタデータは、ファイル処理手段5に入力される。

ここで、前述したように入力手段2では、コンテンツDB1から再生された垂直同期信号に同期して、音声信号をデータ化している。

従って、ファイル処理手段5では、入力手段2からの同期情報と、コンテンツDB1から供給されるタイムコード値とを用いて、音声認識手段3から出力されたメタデータに、file開始からの1秒ごとの基準時間(TM_ENT (秒))と、基準時間からのフレームオフセット数を示す(TM_OFFSET)と、タイムコードを付記した形式でfile化処理する。

例えば、前述した料理番組の場合は、図4に示したようなTEXT形式のメタデータファイルが、ファイル処理手段5から出力される。

次に、記録手段7ではファイル処理手段5から出力されたメタデータファイルとコンテンツDB1から出力されたコンテンツを記録する。

ここで、記録手段7は、HDD、メモリ、光ディスク等から構成されており、コンテンツDB1から出力されたコンテンツもファイル形式で記録する。

【0013】

(実施の形態2)

次に、実施の形態2について説明する。

実施の形態2は、図5に示すように、実施の形態1に対して、コンテンツ情報ファイル処理手段6が付加されている。前記コンテンツ情報ファイル処理手段6では、記録手段7に記録されたコンテンツの記録位置関係を示す制御ファイルを発生し、記録手段7に記録する。

即ち、前記コンテンツ情報ファイル処理手段6では、コンテンツDB1から出力

されたコンテンツと、記録手段7から出力されるコンテンツの記録位置情報をもとに、前記コンテンツが保有している時間軸情報と、記録手段7に記録したコンテンツのアドレス関係を示す情報を発生し、データ化して制御ファイルとして出力する。

例えば、図6に示すように、前記コンテンツの記録位置を示す記録メディアアドレスに対し、前記コンテンツの時間軸基準を示す、 $TM_ENT\ j$ を等時間軸間隔にポイントする。例えば、 $TM_ENT\ j$ を1秒（NTSC信号の場合、30フレーム）毎に記録メディアアドレスをポイントする。

前記のようにマッピングすることで、コンテンツが1秒単位毎に分散記録されても、 $TM_ENT\ j$ から記録手段7の記録アドレスを一義的に求めることができる。さらに、図4で前述したようにメタデータファイルには、ファイル開始からの1秒ごとの基準時間（ TM_ENT （秒））と、基準時間からのフレームオフセット数を示す（ TM_OFFSET ）と、タイムコードと、メタデータとがTEXT形式で記録されている。

従って、前記メタデータファイルの中でメタデータ1を指定すれば、タイムコード、基準時間、及び、フレームオフセット値がわかるので、図6に示す制御ファイルから記録手段7での記録位置が即座にわかることになる。

なお、ここでは前記 $TM_ENT\ j$ の等時間軸間隔は例えば、1秒おきにポイントとした例について説明したが、MPEG2圧縮等で用いられているGOP単位等に合わせて記述することもできる。

さらに、テレビジョン信号のNTSCでは垂直同期信号が60/1.001Hzであるため、絶対時間にあわせるためにドロップフレームモードに合わせたタイムコードと、前記垂直同期信号（60/1.001 Hz）にあわせたノンドロップタイムコードの2種類をしようする。この場合、ノンドロップタイムコードを $TM_ENT\ j$ であらわし、 $TC_ENT\ j$ をドロップフレーム対応タイムコードであらわして使用することもできる。

さらに、制御ファイルのデータ化は、SMIL2等の既存言語を用いてデータ化することも可能であり、さらに、SMIL2の機能をもちいれば、関連したコンテンツ、及び、メタデータファイルのファイル名も合わせてデータ化して、制御ファイル

に格納することができる。

さらに、図6では記録手段の記録アドレスを直接表示する構成をしめしたが、記録アドレスの代わりに、コンテンツファイルの頭からタイムコードまでのデータ容量を表示し、前記データ容量とファイルシステムの記録アドレスから記録手段でのタイムコードの記録アドレスを計算し、検出してもよい。

また、本実施例では、TM_ENT j とタイムコードの対応テーブルをメタデータファイルに格納する形式で説明したが、前記TM_ENT j とタイムコードの対応テーブルは制御ファイル中に格納しても同様の効果がえられる。

(実施の形態3)

次に、実施の形態3について説明する。

【0014】

実施の形態3は、図7に示すように、実施の形態2に対して、検索手段8が付加されている。前記検索手段8では検索したいシーンのキーワードを音声認識してメタデータを検出するのに使用した同一辞書DB4から選択し、設定する。次に、検索手段8では前記メタデータファイルのメタデータ項目をサーチしてキーワードと一致するタイトル名とコンテンツシーンの位置(タイムコード)の一覧を表示する。

【0015】

次に、一覧表示の中から、ひとつの特定シーンが設定された場合は、メタデータファイルの前記基準時間(TM_ENT (秒))と、フレームオフセット数(TM_OFFSET)から制御ファイル中の記録メディアアドレスとを自動的に検出して記録手段7に設定し、前記記録手段7から記録メディアアドレスに記録されたコンテンツシーンをモニタ11に再生する。上記のように構成することで、メタデータを検出して即座に、見たいシーンを検出できる装置を提供できる。

なお、コンテンツにリンクしたサムネイルファイルを準備しておけば、前述したキーワードに一致したコンテンツ名の一覧を表示する際、コンテンツの代表的サムネイル画を再生して表示することも可能である。

(実施の形態4)

次に、他の実施の形態について説明する。

前述の実施形態 1～3 は、あらかじめ記録されているコンテンツにメタデータを付与するシステムについて述べたが、本発明をカメラ等、撮影時にメタデータを付与するシステム、特に、コンテンツ内容が前もって限定される風景撮り、或いは、撮影位置をメタデータとして付与するシステムに対して拡張できる。

このシステムを実施の形態 4 として、図 6 にその構成を示す。

カメラ 51 で撮像し、コンテンツ DB 54 に映像コンテンツが記録されると同時に、カメラが撮影している場所を GPS 52 によって検出し、前記 GPS 52 から出力された位置情報（経緯度数値）を音声合成手段 53 で音声信号化した位置情報も別音声チャンネルに記録する。この場合、記録手段付きカメラ 50 として、カメラ 51、GPS 52、音声合成 53、コンテンツ DB 54 を一体構成してもよい。

次に、コンテンツ DB 54 では前記音声チャンネルに記録されている音声信号の位置情報を音声認識手段 56 に入力する。

ここで、端子 105 から、キー入力ボード等によって、辞書 DB 55 の地域名、ランドマーク等を選択、制限し、前記音声認識手段 56 に出力する。

音声認識手段 56 では認識された経緯度数値と辞書 DB 55 のデータを用いて地域名、ランドマークを検出し、ファイル処理手段 57 に出力する。

次に、ファイル処理手段 57 では、コンテンツ DB 54 から出力されたタイムコードと音声認識手段 56 から出力された地域名、ランドマークをメタデータとして TEXT 化してメタデータファイルを発生させる。

次に、記録手段 58 ではファイル処理手段 57 から出力されたメタデータファイルとコンテンツ DB 54 から出力されたコンテンツデータを記録する。

このように構成することで、撮影したシーン毎に、自動的に地域名、ランドマークのメタデータを付加することができる。

【0016】

なお、上記の各実施形態において、一般的には、音声認識には何らかの誤認識が生じる可能性がある。誤認識が生じた場合、制作されたメタデータ、タグをコンピュータ手段などの情報処理手段を用いて修正することも可能である。

【0017】

また、本発明に係る音声認識手段は単語単位で音声認識する単語認識方式とし、音声入力の単語数、及び、使用する認識辞書の単語数を制限することで、特に、音声認識率を改善することができる。

【0018】

また、本発明では音声認識手段により認識したキーワードをタイムコード共に、メタデータファイルでファイル化する構成を記述したが、音声認識手段により認識したキーワードに加え、関連したキーワードを追加してファイル化してもよい。

【0019】

例えば、音声で淀川を認識した場合は、地形、川等の一般属性キーワードも付加してファイル化する。こうすることで検索時、付加された地形、川等のキーワードも使用することができるので検索性を向上することができる。

【0020】

【発明の効果】

以上説明したように本発明は、コンテンツに関連したメタデータの作成あるいはタグ付けを行うに当たり、制作されたコンテンツのシナリオ等から事前に抽出したキーワードを音声信号として入力し、また、前記シナリオに基づいて辞書分野の設定、及び、キーワードの優先順位つけをおこなっているため、効率よく、正確に音声認識手段からメタデータを発生することができる。

特に、数秒単位の間隔でメタデータを付与する場合は、キー入力では困難であるが、本構成のような音声入力、音声認識を用いれば、数秒単位間隔であっても効率よく、メタデータを付与できる。

また、前記コンテンツの記録位置を示す制御ファイルとメタデータ、及び、タイムコード等を示す前記メタデータファイルとを使用することによって、メタキーワードから一義的に必要なシーンを検索し、前記記録手段から再生することができる。

【図面の簡単な説明】

【図1】

本発明の実施の形態1に係るメタデータ制作装置の構成を示すブロック図

【図 2】

本発明に係る辞書DBの一例を示す構成図

【図 3】

本発明に係るシナリオの一例を示すレシピ図

【図 4】

本発明に係るメタデータファイルの一例を示すTEXT形式のデータ図

【図 5】

本発明の実施の形態 2 に係るメタデータ検索装置の構成を示すブロック図

【図 6】

本発明の情報ファイルの一例を示す構成図

【図 7】

本発明の実施の形態 3 に係るメタデータ検索装置の構成を示すブロック図

【図 8】

本発明の実施形態 4 に係るメタデータ制作装置の構成を示すブロック図

【符号の説明】

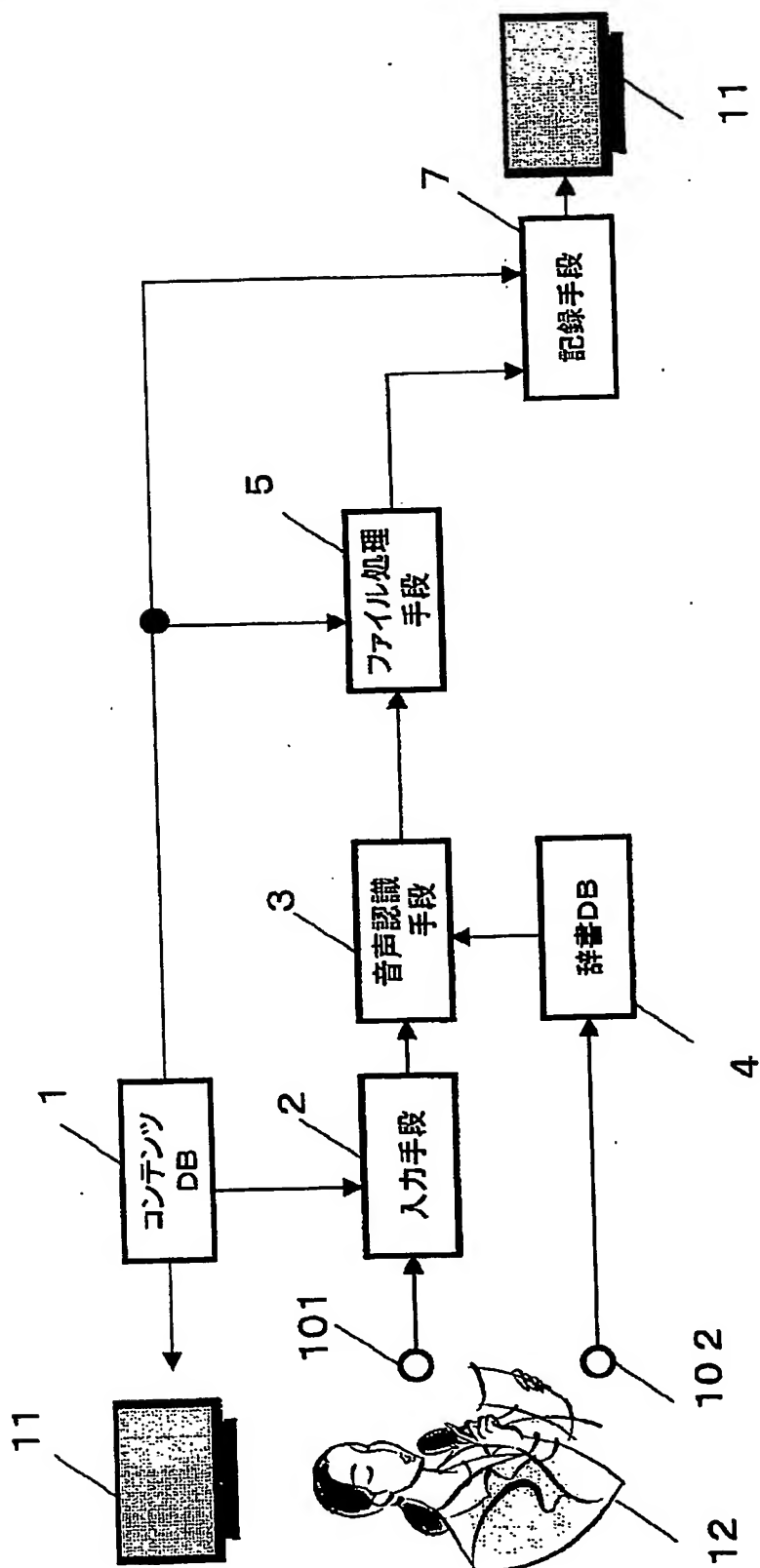
- 1 コンテンツDB
- 2 入力手段
- 3 音声認識手段
- 4 辞書DB 4
- 5 ファイル処理手段
- 6 コンテンツ情報ファイル処理手段
- 7 記録手段
 - 11 映像モニタ
 - 50 記録装置付きカメラ
 - 51 カメラ
 - 52 GPS
 - 53 音声合成手段
 - 54 コンテンツDB
 - 55 辞書DB

- 56 音声認識手段
- 57 ファイル処理手段
- 58 記録手段
- 101 音声入力端子
- 102 辞書分野選択入力端子
- 105 辞書地名選択入力端子

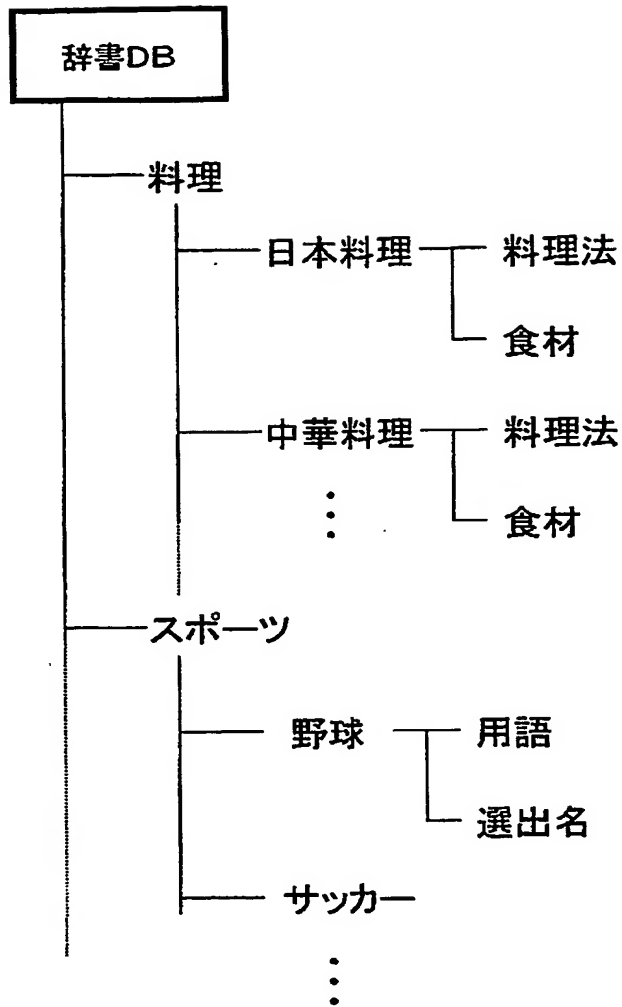
【書類名】

図面

【図 1】



【図 2】



【図3】

レシピ

<メニュー>アサリのトマト煮 白ワイン風味

<材料>4人分

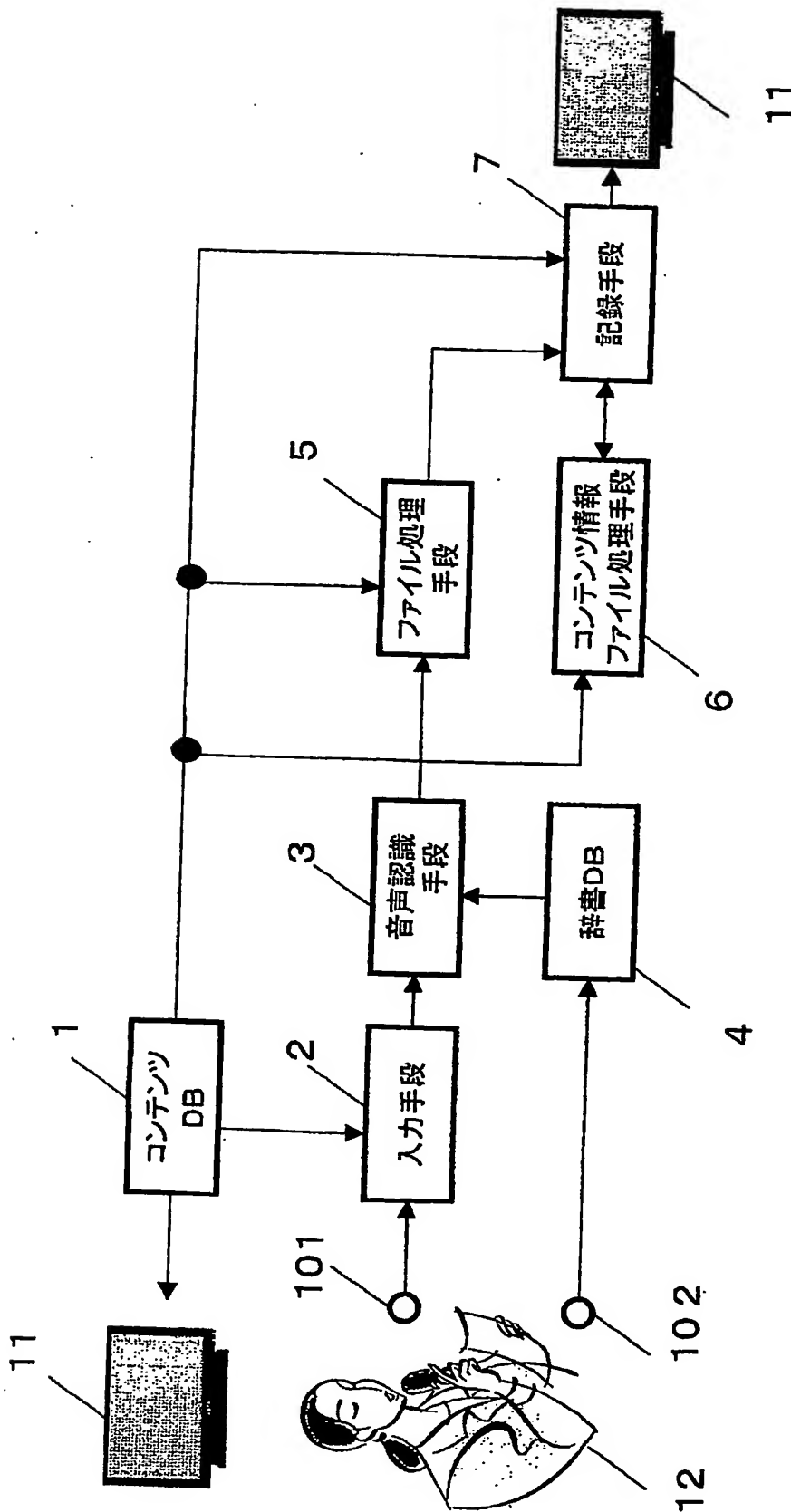
材 料	分量
あさり	50個
カキ	10個
玉ねぎ(小)	1個分
トマト(湯むきし小さくきざんだもの)	1個分
にんにく(みじん切り)	2枝
白ワイン	1カップ
タイム	少々
パセリ(みじん切り)	1/2カップ
オリーブ油	1/4カップ

【図 4】

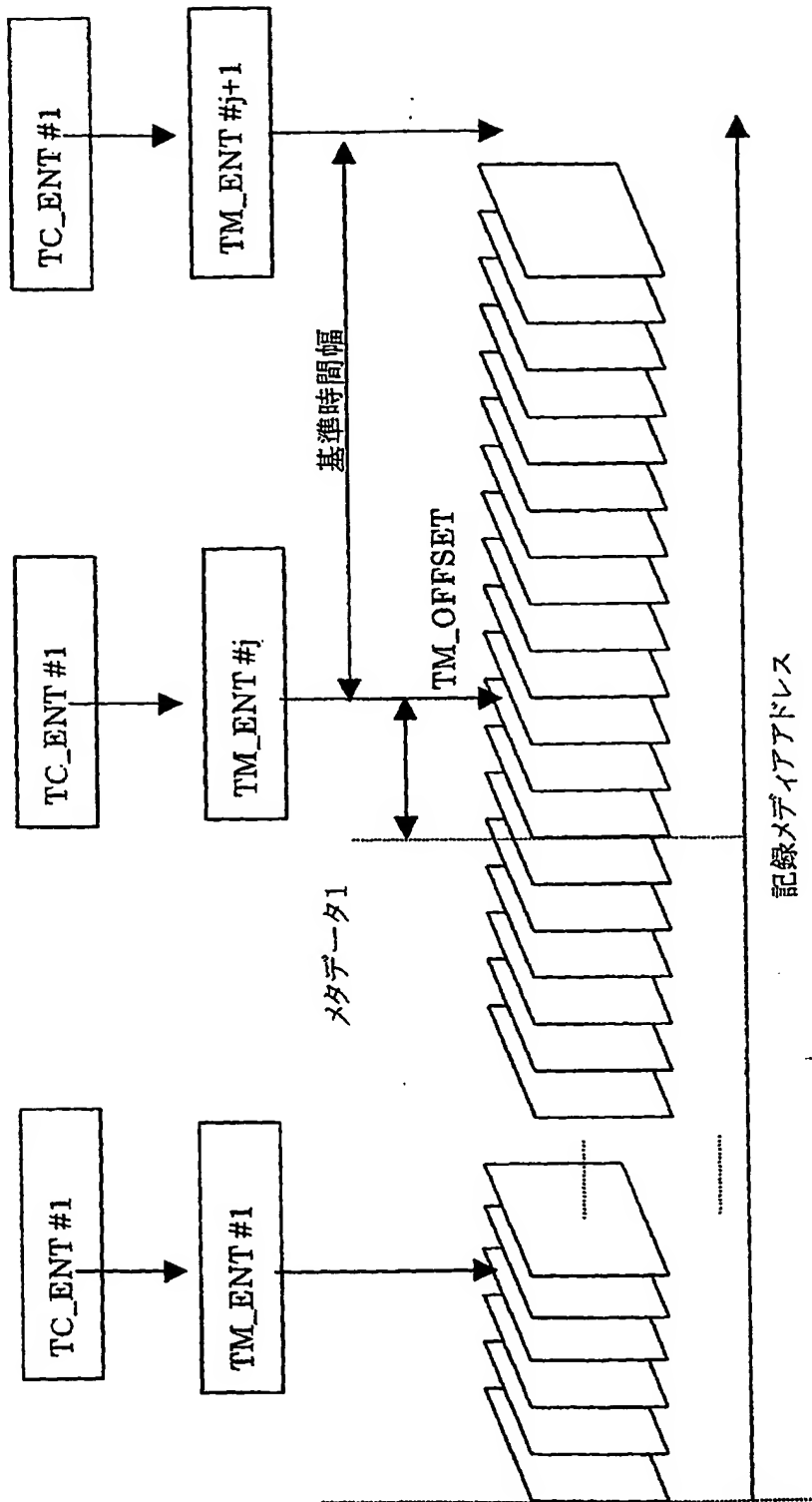
Title: あさりのトマト煮			
TM_ENT #j	TM_OFFSET*1	TC	メタデータ
-----	-----	-----	-----
.	.	.	.
930	+8	00 15 30 08	玉ねぎ 1個
990	+5	00 16 30 05	にんにく 2枝
.	.	.	.
.	.	.	.
1230	+10	00 20 30 10	あさり、カキ
.	.	.	炒める
.	.	.	.

*1: オフセットフレーム数で表記

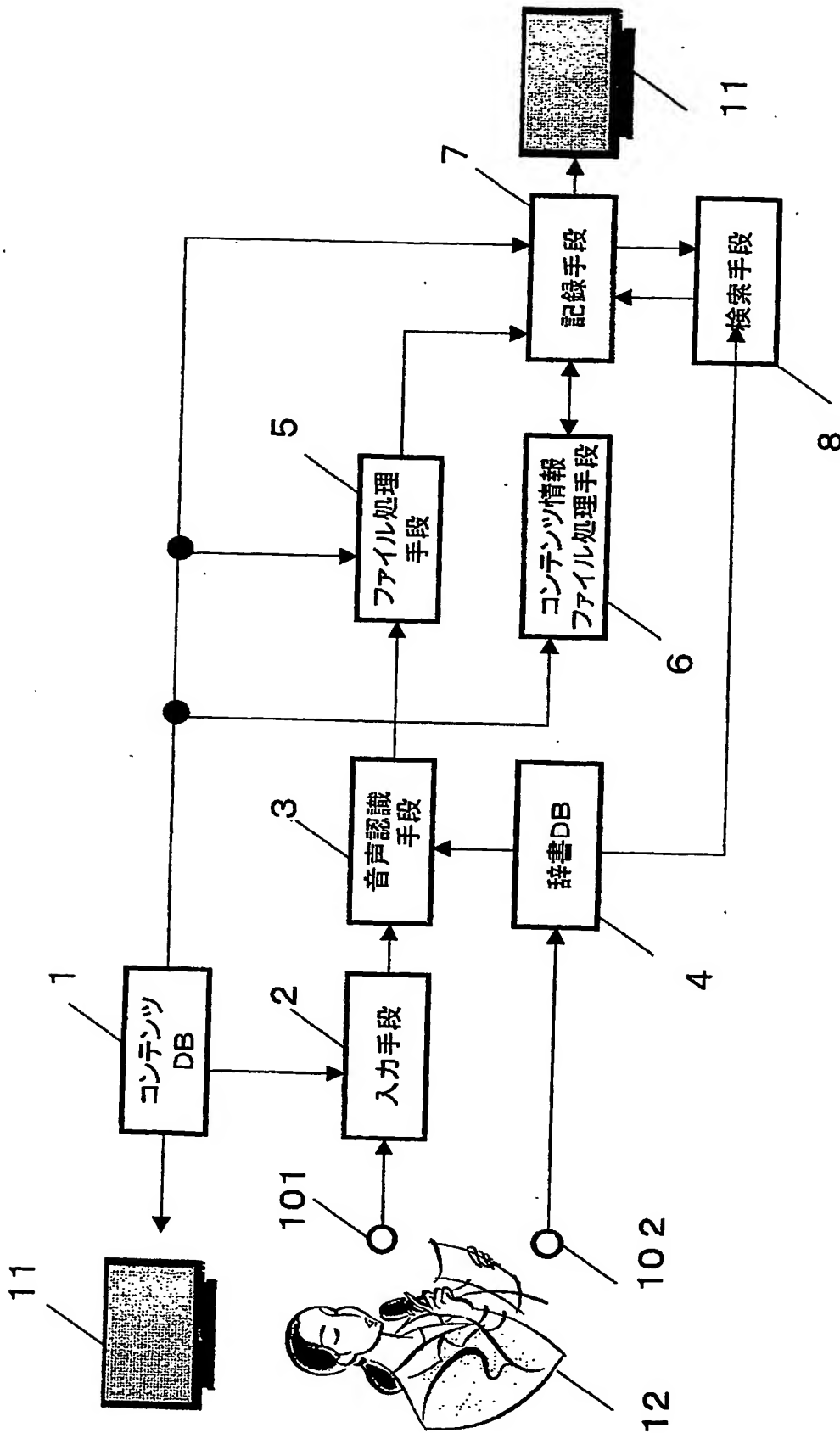
【図 5】



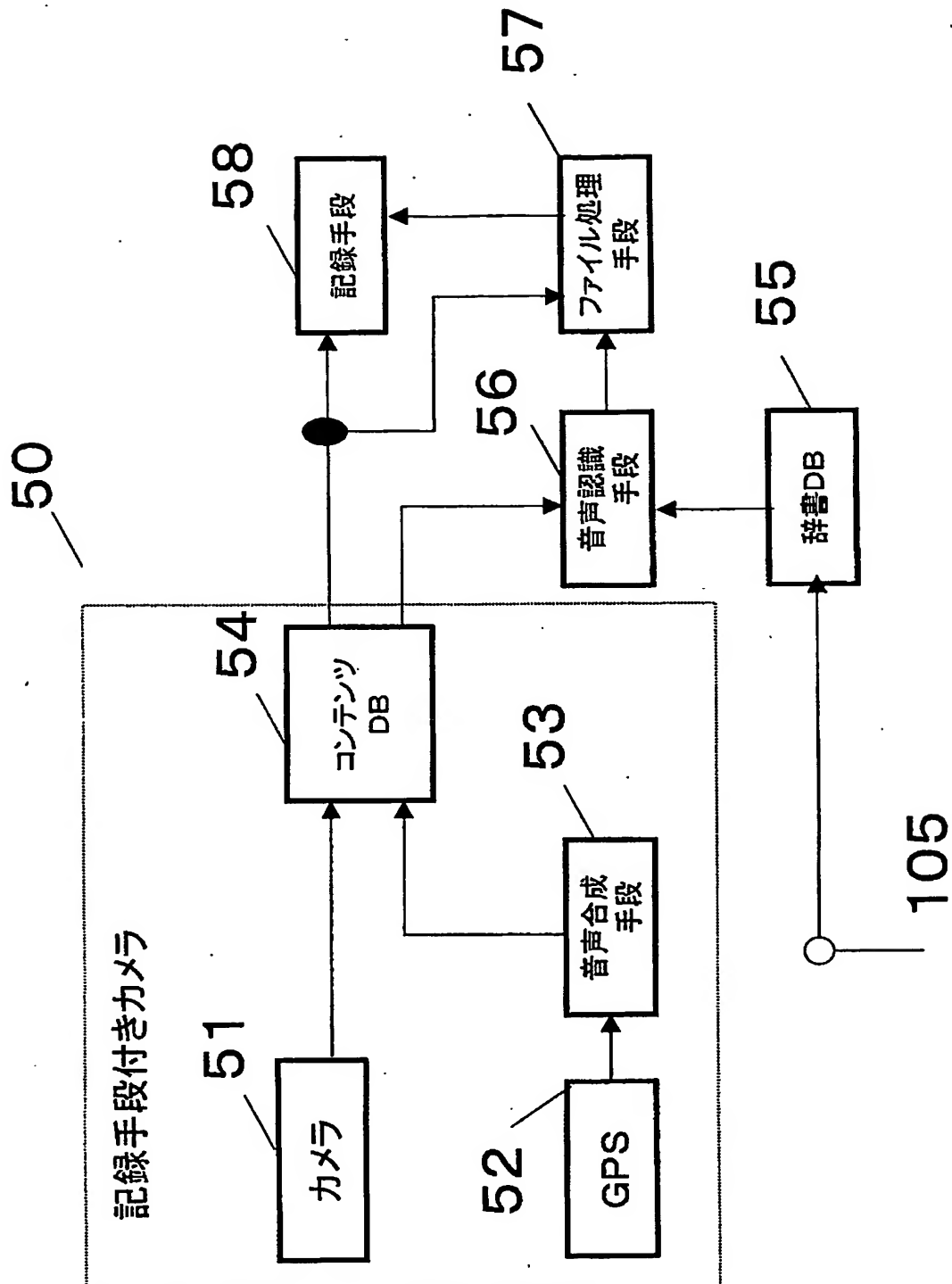
【図6】



【図7】



【図 8】



【書類名】 要約書

【要約】

【課題】 制作された映像・音声コンテンツを再生することによりメタデータとすべき情報を確認し、音声入力でコンピュータ等に入力することにより前記メタデータを制作し、検索するシステムを提供することを目的とする。

【課題解決】 コンテンツに関連するメタデータの制作装置であって、制作されたコンテンツに合わせて制作されたシナリオ、或いは前記コンテンツの内容から抽出されたコンテンツ管理用キーワードの音声信号を入力し、前記音声信号をデータ化する入力手段と、前記入力手段でデータ化された音声信号データから、キーワードを認識する音声認識手段と、前記音声認識手段から出力されたキーワードを、コンテンツに含まれる画像信号との時間位置を示すタイムコードと共にメタデータファイルに記憶するファイル処理手段とを具備する。

【選択図】 図 1

特願 2002-319756

出願人履歴情報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社